Data Analytics In Preventive Health Care With Total Quality Management Approach

Özlem Akcay Kasapoğlu, PhD
Istanbul University, Turkey

Abstract

In this new era where people are more concerned about their health, they are trying to follow a healthy life style. The hospital and health care units is continually improving their value to their patients. This they accomplish through the provision of high quality services. Since there is a worldwide rise in the prevalence of chronic diseases, preventive healthcare should be one of the services that should be given to patients in the customer relationship management concept. This service could be integrated to the hospitals and local healthcare units with an understanding of total quality management and with a strong top management support. Being the customer's focus, understanding the customer is an important total quality management principle. Health care is the most crucial and delicate service that could be given. With preventive health care practices, effective patient data tracking, and data analytics, patients with potential diseases could be identified. Also, better life choice practices could be recommended which would prolong their life expectancy. Health promotional activities do not target a specific disease or condition, but rather promote the health and well-being on a general level. In this article, with the help of data mining tools, the patients' potential risks were identified. In addition, different health promotion and training services were recommended to be given to the patients. In training the patients, the effects of life choices from nutritious meals to daily exercising, increase the overall well besides preventing some of the potential threats. Data mining methods were seen as a useful supplementary method in analysing preventive health care data and its capability to illustrate a large dataset. Therefore, its relationships between variables are the reason it was selected in this study.

Keywords: TQM, Preventive Health Care, Customer Satisfaction, Data Mining, CRM

Introduction

Preventive healthcare consists of measures taken for disease prevention. Thus, this is as opposed to the treatement of disease. Just as health encompasses a variety of physical and mental states, so do disease and disability. They are affected by environmental factors, genetic predisposition, disease agents, and lifestyle choices. Health, disease, and disability are dynamic processes which begin before individuals realize that they are affected. Disease prevention relies on anticipatory actions that can be categorized as primary, secondary, and tertiary prevention.

Each year, millions of people die of preventable deaths. Consequently, there are several methods used for the prevention of disease. It is recommended that adults and children should visit their doctor for regular check-ups, even if they feel healthy. Furthermore, they were asked to perform disease screening, identify risk factors for disease, discuss tips for a healthy and balanced lifestyle, stay up to date with immunizations and boosters, and maintain a good relationship with a healthcare provider. Some common disease screenings include checking for high blood pressure, high blood sugar (a risk factor for diabetes), high blood cholesterol, screening for colon cancer, and depression. In addition, these diseases include HIV and other common types of sexually transmitted disease such as chlamydia, syphilis, gonorrhea, mammography (to screen for breast cancer), colorectal cancer screening, a pap test (to check for cervical cancer), and screening for osteoporosis (Preventive Health Care, 2016).

Health care as a service provider is a very delicate and crucial topic considering the importance of the human life. In all of the patient’s services, zero defect must be considered. Zero defect concept is not an easy goal to succeed. Continuing improvement philosophy must be learned, adapted, and digested. Therefore, this philosophy is a very crucial part of the total quality management practices. If a health care service provider unit from a small clinic to a big hospital can be adapted to total quality management, then an understanding of continuous improvement comes along with it. Subsequently, this can only succeed with the top management’s leadership and support. Like the concept of continuous improvement, being the customer’s focus is part of TQM approach in the health care system. Therefore, it should be corrected to be more focused on patients.

According to the ASQ (2016), a core definition of total quality management (TQM) describes a management approach to long–term success through customer satisfaction. With the use of TQM, all members of an organization participate in improving processes, products, services, and the culture in which they work. Total quality management can be summarized as a management system for a customer-focused organization that involves all employees in continual improvement. Also, it uses strategy, data, and effective communications to integrate the quality discipline into the culture and activities of the organization. The customer ultimately determines the level of quality. However, no matter what an organization does to foster quality improvement—training employees, integrating quality into the design process, upgrading computers or software, or buying new measuring tools—the customer determines whether the efforts is worthwhile.

Furthermore, while customer relationship management is a core approach in managing interactions between commercial organizations, typically, banks and retailers and their customers is not less important in a healthcare context. Customer interactions may occur through call centers, physicians' offices, billing departments, inpatient settings, and ambulatory care settings. As in the case of commercial organizations, data mining applications can be developed in the healthcare industry to determine the references, usage patterns, and the current and future needs of individuals. This is employed for the purpose of improving their level of satisfaction. Therefore, these applications can also be used to predict other products that a healthcare customer is likely to purchase. This involves whether a patient is likely to comply with prescribed treatment or whether preventive care is likely to produce a significant reduction in future utilization. To aid healthcare management, data mining applications can be developed to better identify and track chronic disease states and high-risk patients, design appropriate interventions, and reduce the number of hospital admissions and claims (Koh & Tan, 2005).

TQM practices plays a very important role in hospitals competitive strategy. Hospitals differentiate themselves from their competitors with their high quality and variety in their service products mix. Health care services does not just include diagnosis, surgery, treatment and medicines, but also preventive health care. Tracking the patient's data and applying a preventive health care strategy should be part of their strategy. In this article, data from the patients are analyzed in order to identify the potential risk carriers. Subsequently, before the diseases occurs, the patients tracked were supposed to be informed about a better health practices. However, this would decrease the diseases and increase customer's satisfaction.

Furthermore, data mining is used for finding previously unknown patterns and trends in databases. Also, it uses this information to build predictive models. In healthcare, data mining is becoming increasingly popular, if not increasingly essential. Also, several factors have motivated the use of data mining applications in healthcare services. The existence of medical insurance, fraud, and abuse, for example, has led many healthcare insurers to attempt to data mining which is considered as a relatively recently developed methodology and technology. Thus, this technology became prominent only in 1994 (Koh & Tan, 2005). In this article, data from both male and female patients with their cholesterol and blood pressure measures were taken. The data also includes those that engages in sports, drink alcohol, and smokers. Also, one of the data mining tools, decision trees, was applied to the data.

Literature Review

In literatures in Total Quality Management and Continuous Improvement, understanding have been gotten from some researches conducted. Valmohammadi & Roshanzamir (2015) worked on the guidelines of improvement. Thus, they searched the relations among organizational culture. Steele & Schomer (2009) studied continuous quality improvement programs which provide new opportunities to drive value innovation initiatives in hospital based radiology practices. Jung & Wang (2006) have studied the relationship between TQM and continuous improvement. Katzman & Paushter (2016) tried to build a culture of continuous quality improvement in an academic radiology department. Hohan, Olaru & Pirnea (2015) studied the assessment and continuous improvement of information security based on TQM and business excellence principles. Consequently, Taylor & Wright (2004) studied the contribution of measurement and information infrastructure to the success of TQM. Furthermore, Taylor (1998) made a TQM implementation on organisation practices and information.The results underline the central role of data and information, and hence measurement, in successful TQM implementations. Lorente, Rodríguez, & Dewhurst (2004) worked on the effect of information technologies on TQM. Also, Daghfous & Barkhi (2009) studied the strategic management of information technology in UAE hotels. They did an exploratory study of TQM, SCM, and CRM implementations. Also, Taylor and Wright (2006) worked on the contribution of measurement and information infrastructure to TQM success.

However, there are studies on information gathering from preventive health care systems. In addition, Fields (1980) worked on client tracking. She developed a state-wide hypertension information system. Shih et al. (2011) worked on the health information systems. In small practices, they improved the delivery of clinical preventive services. Despite strong evidence that clinical preventive services (cps) reduce morbidity and mortality, cps performance has not improved in adult primary care. Redmond et al. (2010) investigated the sources of health information related to preventive health behaviors in a national study. Current literature suggests that certain sources of information are used in varying degrees among different socioeconomic and demographic groups. Therefore, it is important to determine if specific classes of health information sources are more effective than others in promoting health behaviors. Their study aims to determine if interpersonal versus mass media sources of health information are associated with meeting recommendations for health behaviors (nonsmoking, fruit/vegetable intake, and exercise) and cancer screening.

More researches are done on data mining and preventive health care. Subsequently, Chang, Wang & Jiang (2011) used data mining techniques for multi-diseases prediction modeling of hypertension and hyperlipidemia by common risk factors. They mentioned that many previous studies have employed predictive models for a specific disease. Therefore, they failed to note that humans often suffer from not only one disease, but from associated diseases as well. Thus, these associated multiple diseases might have reciprocal effects. Also, abnormalities in physiological indicators can indicate multiple associated diseases. In addition,, common risk factors can be used to predict the multiple associated diseases. However, their approach provides a more effective and comprehensive forecasting mechanism for preventive medicine. Nenonen (2013) analysed various factors related to slipping, stumbling, and falling accidents at work. The utilisation of data mining methods has become common in many fields. In occupational accident analysis, however, these methods are still rarely exploited. Their study applies methods of data mining (decision tree and association rules) to the Finnish national occupational accidents and diseases statistics database. Cheng, Yao & Wu (2013) applied data mining techniques to analyze the causes of major occupational accidents in the petrochemical industry. Accidents that occur in the petrochemical industry frequently result in serious social issues. Behind every occupational accident, there are safety management problems requiring investigation. This study collected 349 cases of major occupational accidents in the petrochemical industry between 2000 and 2010 in Taiwan for analysis. Using descriptive statistics, they elucidated the factor distribution of these major occupational accidents. The Classification And Regression Tree (CART) was used to examine the distribution and rules governing the factors of the disasters. Mookiah et al. (2012) used data mining technique for automated diagnosis of glaucoma using higher order spectra and wavelet energy features. Eye images provide an insight into important parts of the visual system. Thus, it also indicate the health of the entire human body. Glaucoma is one of the most common causes of blindness. It is a disease in which fluid pressure in the eye increases gradually, damages the optic nerve, and result to the loss of vision. Robust mass screening may help to extend the symptom-free life for the affected patients. Shenas et al. (2014) identified high-cost patients using data mining techniques and a small set of non-trivial attributes. In their paper, they used data mining techniques, namely neural networks and decision trees, to build predictive models. This is used for the purpose of identifying very high-cost patients in the top 5 percentile among the general population.

They mentioned that the results of this study can be used by healthcare data analysts, policy makers, insurer, and healthcare planners to improve the delivery of health services.

Hajakbari & Minaei-Bidgoli (2014) found a new scoring system for assessing the risk of occupational accidents. However, they performed a case study using data mining techniques with Iran's Ministry of Labor data. Yeh, Cheng & Chen (2011) constructed a predictive model for cerebrovascular disease using data mining. Also, Cheng et al. (2012) applied data mining techniques to explore factors contributing to occupational injuries in Taiwan's construction industry they have worked on. They did accident analysis and prevention. Raju et al. (2015) explored factors associated with pressure ulcers: with a data mining approach.

Lee, Chen & Tseng (2011) worked on a novel data mining mechanism considering bio-signal and environmental data with applications on asthma monitoring. Silva & Jacinto (2012) tried to find occupational accident patterns in the extractive industry using a systematic data mining approach in the Portuguese Extractive Industry. Su et al. (2006) used data mining for the diagnosis of type II diabetes from three-dimensional body surface anthropometrical scanning data. Consequently, Yeh, Wu & Tsao (2011) used data mining techniques to predict hospitalization of hemodialysis patients. This study combines temporal abstraction with data mining techniques for analyzing dialysis patients' biochemical data in developing a decision support system. The mined temporal patterns are helpful for clinicians to predict hospitalization of hemodialysis patients and to suggest immediate treatments to avoid hospitalization. Vijayakrishnan et al. (2014) studied the prevalence of heart failure signs and symptoms in a large primary care population. Hence, this was identified through the use of text and data mining of the electronic health record. Chen, Chouq & Hwang (2003) made an application of a data-mining technique to analyze coprescription patterns for antacids in Taiwan. Therefore, the aim of this study is to estimate the scale of antacid prescription in Taiwan. This is done using the national insurance claims for outpatient services. It also aim to analyze coprescribing patterns of antacids using modern data-mining techniques. A data mining technique -association rule mining- was applied to identify the drugs prescribed in combination with antacids. Suzuki et al. (2015) worked on the comedications alter drug-induced liver injury reporting frequency with data mining. They examined the effect of these drug–drug interactions on liver safety reports of four drugs that is highly associated with hepatotoxicity. Co-reported medications were associated with changes in the liver event reporting frequency of drugs commonly associated with hepatotoxicity. Also, it suggest that comedications may modify drug hepatic safety. Chen et al. (2016) performed personal health indexing based on medical examinations with a data mining approach. Kalaitzopoulos, Patel & Younesi (2016) worked on the advancements, in data management and data mining approaches, in translational medicine.

Data Mining

Data mining aims to identify valid, novel, potentially useful, and understandable correlations and patterns in data. This is done by combing it through copious data sets to sniff out patterns that are too subtle or complex for humans to detect. Cross-Industry Standard Process for Data Mining, or CRISP-DM (see www.crisp-dm.org) proposes the following methodology for data mining: business understanding, data understanding and preparation, modeling, evaluation, and deployment. Business understanding is critical because it identifies the business objectives and, thus, the success criteria of data mining projects. Furthermore, as the term “data mining” implies, data is a crucial component. Therefore, no data means no mining. Hence, CRISP-DM includes data understanding and data preparation. In other words, sampling and data transformation are essential antecedents for modeling. The modeling stage is the actual data analysis. Most data mining software include online analytical processing; traditional statistical methods, such as cluster analysis, discriminant analysis, and regression analysis; and non-traditional statistical analysis, such as neural networks, decision trees, link analysis, and association analysis. Consequently, this extensive range of techniques is not surprising in the light of the fact that data mining has been viewed as the offspring of three different disciplines. These discipline are database management, statistics, and computer science, including artificial intelligence and machine learning. The evaluation stage enables the comparison of models and results from any data mining model by using a common yardstick, such as lift charts, profit charts, or diagnostic classification charts. Finally, deployment relates to the actual implementation and operationalization of the data mining models. Data mining techniques can be broadly classified based on what they can do. They include description and visualization; association and clustering; and classification and estimation, which is predictive modeling. Description and visualization can contribute greatly towards understanding a data set, especially a large one. It can also be used in detecting hidden patterns in data, especially complicated data containing complex and non-linear interactions. They are usually performed before modeling is attempted. Thus, they represent data understanding in the CRISP-DM methodology. Based on association, the objective is to determine which variables go together; for example, market-basket analysis (Koh & Tan, 2005).

Nowadays, each individual and organization can access a large quantity of data and information about itself and its environment. This data has the potential to predict the evolution of interesting variables or trends in the outside environment. Nevertheless, there are two main problems of data. Firstly, information is scattered within different archive systems that are not connected with one another, thereby producing an inefficient organization of the data. Secondly, there is a lack of awareness about new analysis tools. Two developments could help in overcoming these problems. First, software and hardware offer power at lower cost which allows organizations to collect and organize data in structures that give easier access and transfer. Second, methodological research, particularly in the field of computing and statistics, has recently led to the development of flexible and scalable procedures that can be used to analyze large data stores. Consequently, these two developments have shown that data mining is rapidly spreading through many businesses as an important intelligence tool for backing up decisions (Giudici, p.1). Data mining is used to refer to a very interdisciplinary field, which consists of using methods of several research areas to extract knowledge from real-world datasets (Freitas, p.1). The knowledge must be new, not obvious; and people or organizations must be able to use it (Adriaans & Zantinge, p.5). Data mining is the practice of automatically searching large stores of data to discover patterns and trends that go beyond simple analysis. Data mining uses sophisticated mathematical algorithms to segment the data and evaluate the probability of future events (Oracle). There are several data mining tasks like classification, clustering, dependence modeling, and association rules. Each task can be considered as a kind of problem to be solved by a data mining algorithm. Therefore, the first step in the development of a data mining algorithm is to define which data mining task the algorithm addresses (Freitas, p.13). In this paper, decision trees algorithm for classification was used.

Decision Trees

Databases are rich with hidden information that can be used for intelligent decision making. Classification and prediction are two forms of data analysis that can be used to extract models describing important data classes or to predict future data trends. Many classification and prediction methods have been proposed by researchers in machine learning, pattern recognition, and statistics. Most algorithms are memory resident, typically assuming a small data size. Recent data mining research has built on such work. They develop scalable classification and prediction techniques capable of handling large disk- resident data. Classification and prediction have numerous applications including fraud detection, target marketing, performance prediction, manufacturing, and medical diagnosis (Han & Kamber, p.285). Decision tree is one of the most widely used techniques for classification which has the task of assigning objects to one of several predefined categories. In practice, one wants to have a small and accurate tree for many reasons. A smaller tree is more general and also tends to be more accurate. It is also easier to understand by human users. In many applications, the user's understanding of the classifier is important (Liu, p. 59-60). The Decision Tree procedure creates a tree-based classification model. It classifies cases into groups or predicts values of a dependent (target) variable based on values of independent (predictor) variables. The procedure provides validation tools for exploratory and confirmatory classification analysis. Therefore, the procedure can be used for:

Segmentation: Identify persons who are likely to be members of a particular group.

Stratification: Assign cases into one of several categories, such as high-, medium-, and low-risk groups.

Prediction: Create rules and use them to predict future events, such as the likelihood that someone will default on a loan or the potential resale value of a vehicle or home.

Data Reduction and Variable Screening: Select a useful subset of predictors from a large set of variables for use in building a formal parametric model.

Interaction Identification: Identify relationships that pertain only to specific subgroups and specify these in a formal parametric model.

Category Merging and Discretizing Continuous Variables: Recode group predictor categories and continuous variables with minimal loss of information.

Subsequently, there are exponentially many decision trees that can be constructed from a given set of attributes. While some of the trees are more accurate than others, finding the optimal tree is computationally infeasible because of the exponential size of the search space. Nevertheless, efficient algorithms have been developed to induce a reasonably accurate, albeit suboptimal, decision tree in a reasonable amount of time. Thus, these algorithms usually employ a greedy strategy that grows a decision tree by making a series of locally optimum decisions about which attribute to use for partitioning the data. ID3, C4.5, C&RT, QUEST are examples of the decision tree algorithms (Tan, Steinbach & Kumar, p.151-152).

The available growing methods is the Chi-squared Automatic Interaction Detection (CHAID). At each step, CHAID chooses the independent (predictor) variable that has the strongest interaction with the dependent variable. Categories of each predictor are merged if they are not significantly different with respect to the dependent variable.

Classification and Regression Trees (C&RT): C&RT splits the data into segments that are as homogeneous as possible with respect to the dependent variable. A terminal node in which all cases have the same value for the dependent variable is known as an homogeneous "pure" node.

Quick, Unbiased, Efficient Statistical Tree (QUEST): A method that is fast and avoids other methods' bias in favor of predictors with many categories. Therefore, QUEST can be specified only if the dependent variable is nominal (IBM, p.1).

Application

Method: In this paper, data mining tools and C&RT was used. Analysis are done in SPSS Clementine. Decision trees are used for classification.

Data Set

A total of 1000 male and female patient's blood pressure and total cholesterol measurements were taken. Blood pressure and total cholesterol values are coded as low, normal, and high. Also, in the data set, there are records of the patients involved in sports, smoking, and who are drinking alcohol. Thus, this made it 7 fields and 1000 records in the dataset. Variables are "age, gender, blood pressure, cholesterol, sport, smoking, and alcohol". The built model is seen in Figure 1. Also, the summary of the variables can be seen in Figure 2.

Figure 1. Model

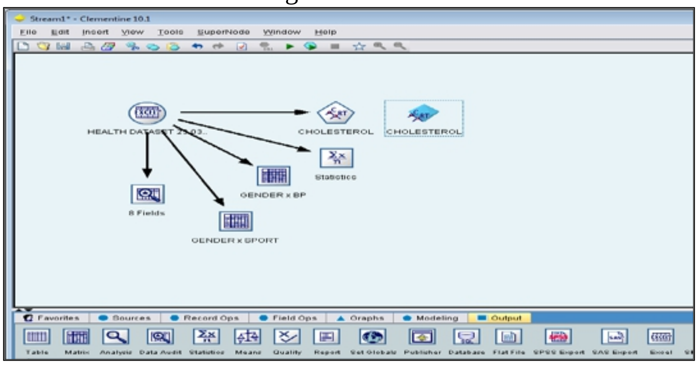
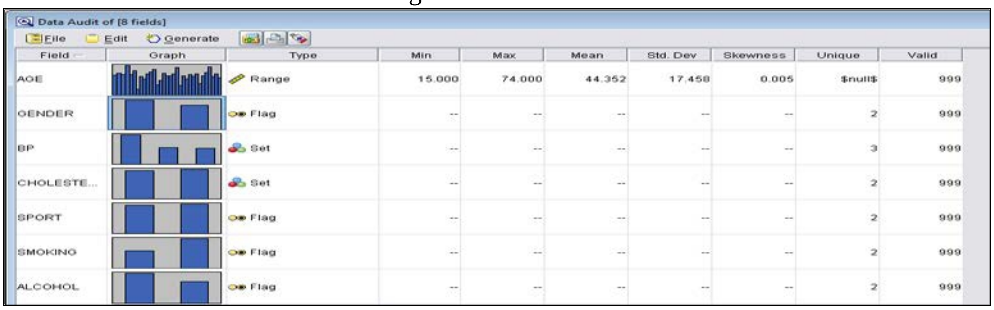


Figure 2. Data Audit



Field	Graph	Type	Min	Max	Mean	Std. Dev	Skewness	Unique	Valid
AGE		Range	15.000	74.000	44.352	17.458	0.005	$null$	999
GENDER		Flag	--	--	--	--	--	2	999
BP		Set	--	--	--	--	--	3	999
CHOLESTE...		Set	--	--	--	--	--	2	999
SPORT		Flag	--	--	--	--	--	2	999
SMOKING		Flag	--	--	--	--	--	2	999
ALCOHOL		Flag	--	--	--	--	--	2	999

Figure 3. Histogram of Age

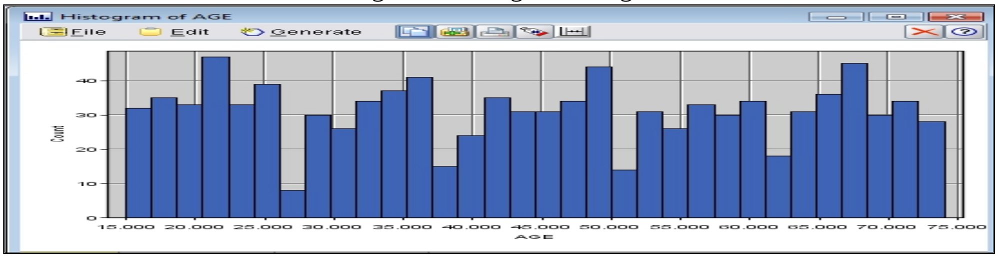
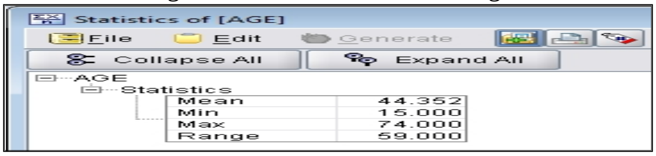


Figure 4.The Statistics for the Age



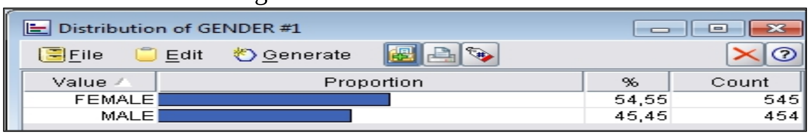
Statistics of [AGE]

AGE
Statistics

Mean	44.352
Min	15.000
Max	74.000
Range	59.000

The mean of the age is 44,352.(Figure 3 and Figure 4). According to Figure 5, 54.55% of the patients are female and 45.45% are male.

Figure 5. Distribution of Gender

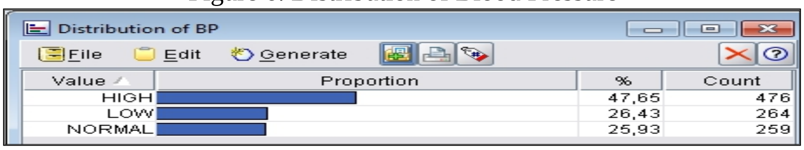


Distribution of GENDER #1

Value	Proportion	%	Count
FEMALE		54,55	545
MALE		45,45	454

47.65% of the patients have high blood pressure, 26.43% of them have low BP, and 25.93% of the patients have normal BP (Figure 6).

Figure 6. Distribution of Blood Pressure

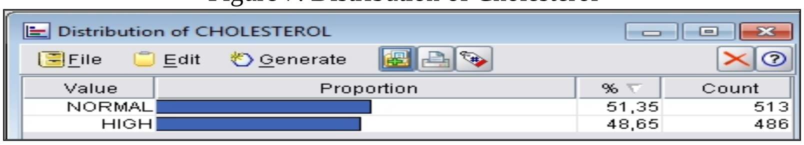


Distribution of BP

Value	Proportion	%	Count
HIGH		47,65	476
LOW		26,43	264
NORMAL		25,93	259

In Figure 7, the distribution of cholesterol can be seen. 51.35% of the patients have normal cholesterol and the rest have high cholesterol.

Figure 7. Distribution of Cholesterol

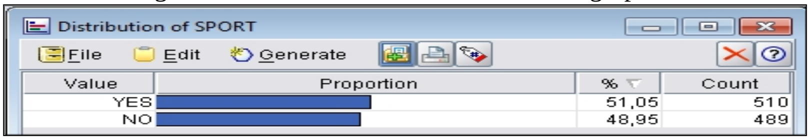


Distribution of CHOLESTEROL

Value	Proportion	%	Count
NORMAL		51,35	513
HIGH		48,65	486

51.05% of the patients engaged in sports.

Figure 8. Distribution of the Patients Doing Sports

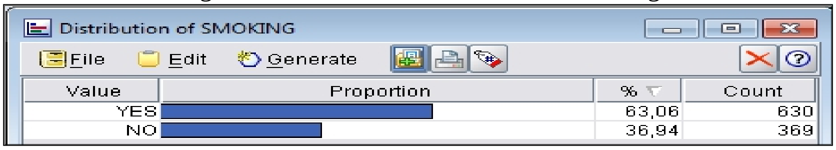


Distribution of SPORT

Value	Proportion	%	Count
YES		51,05	510
NO		48,95	489

63.06% of the patients smoke (Figure 9).

Figure 9. Distribution of the Patient Smoking

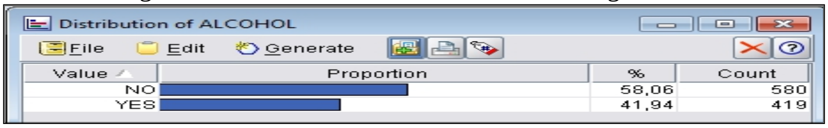


Distribution of SMOKING

Value	Proportion	%	Count
YES		63,06	630
NO		36,94	369

According to Figure 10, 58.06% of the patients do not drink alcohol.

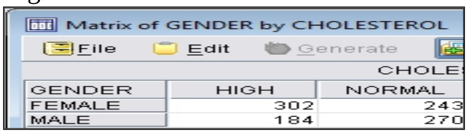
Figure 10. Distribution of the Patient Drinking Alcohol



Distribution of ALCOHOL

Value	Proportion	%	Count
NO		58,06	580
YES		41,94	419

Figure 11. Matrix of Gender to Cholesterol

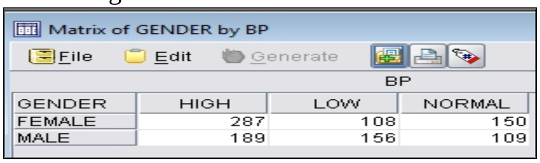


Matrix of GENDER by CHOLESTEROL

GENDER	HIGH	NORMAL
FEMALE	302	243
MALE	184	270

In Figure 11, the matrix of gender to cholesterol shows that 62% of the female have high cholestrerol (302 female), while 52% of the male have normal cholesterol (270 male).

Figure 12. Matrix of Gender to BP

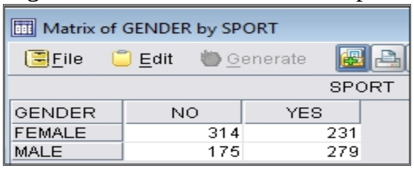


Matrix of GENDER by BP

GENDER	HIGH	LOW	NORMAL
FEMALE	287	108	150
MALE	189	156	109

In the Figure 12, it can be seen that 60% of the females' blood pressure is high (287 female). 59% of the male have low blood pressure (156 male) and 58% of the female have normal blood pressure (150 female). In Figure 13, one can see that 64% of the female do not engage in sport (314 female), while 45% of the male do engage in sport (279 male).

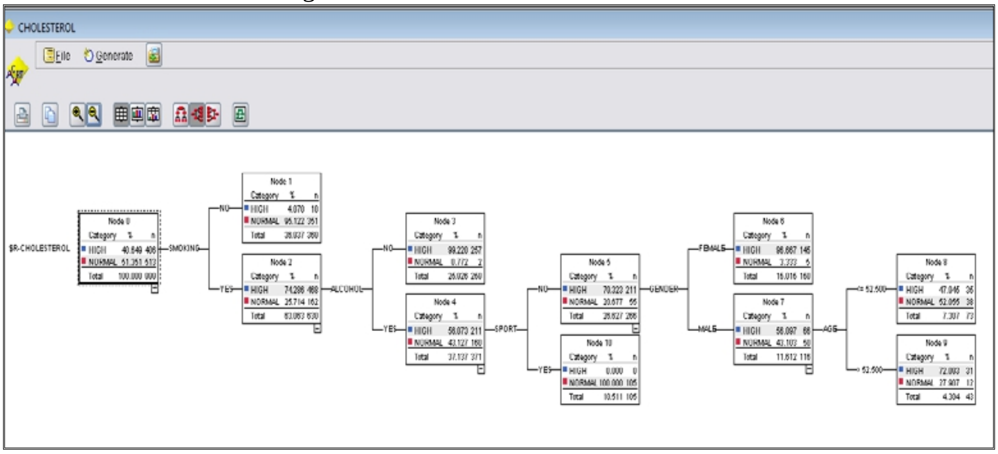
Figure 13. Matrix of Gender to Sports



Matrix of GENDER by SPORT

	SPORT	
GENDER	NO	YES
FEMALE	314	231
MALE	175	279

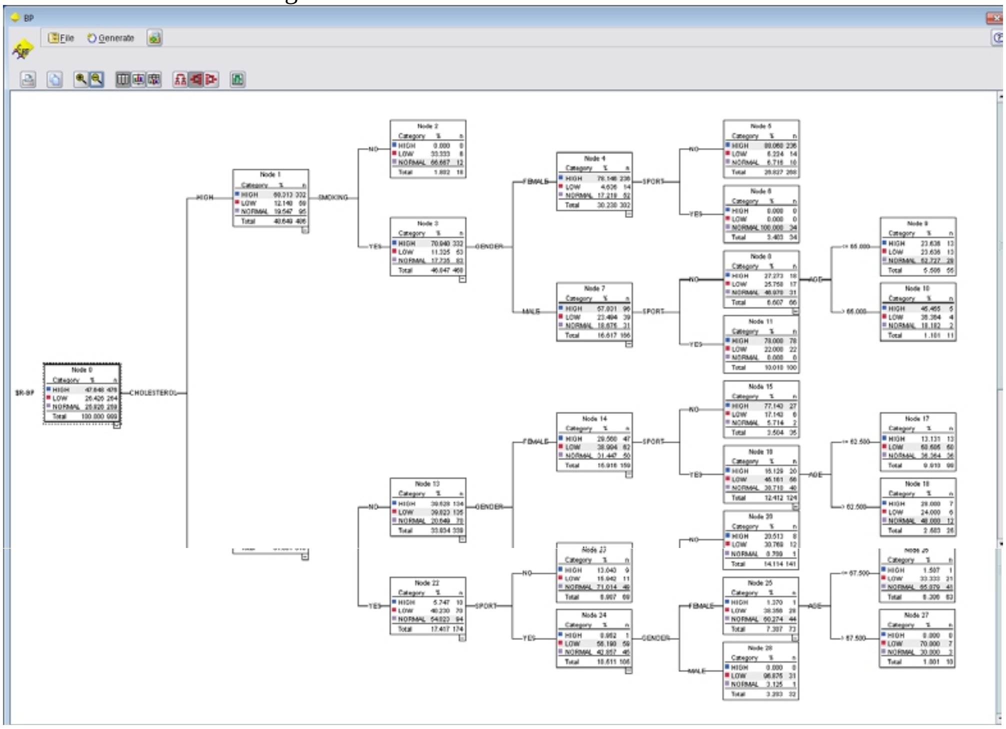
Figure 14. Decision Tree for Cholesterol



In Figure 14, the analysis done for decison trees shows that 48.649% of the patients have high cholesterol. 51.351% of the patients have low cholesterol. Smoking variable is effective here and it can be read from the tree that 95.122% of the non smokers have abnormal cholesterol. Furthermore, 51.351% ratio patients have high and low cholestreol, while 95.122% of them are non smokers.

74.286% of the smokers have high cholesterol if they are not drinking alcohol. 99.228% of them have high cholesterol. Thus, if they are drinking alcohol, 56.873% of them have high cholesterol. Engaging in sports on a regular basis is important. In our data, it was found that out of all the people who are smokers and are drinking alcohol, those engaged in sports have a normal cholesterol of 100%. On the other hand, the ones who are not engaged in sports have a high cholesterol with 79.33%. 96.667% of the females who are smokers, drinking alcohol, and not engaging in sports have high cholesterol, while the 56.897% of male have high cholesterol. In the same branch, age is also effective. The male who have an age under 52.5 with 52.055 % have normal cholesterol, while the ones who are above the 52.5 age with 72.093% have high cholesterol.

Figure 15. Decision Tree for Blood Pressure



In the blood pressure, cholesterol is effective. 68.313% of the patients who have high cholesterols also have high blood pressure. 88.060% of the patients who has high cholesterol, smokes, and who are not engaging in sports have high blood pressure. 100% of the female who has a high cholesterol, smoking, and engaging in sports have normal blood pressure. 27.273% of the male who has high cholesterol, smoking, and not engaging in sports have high blood pressure. In this branch, age is effective. However, 45.455% of them who have an age older than 65 have high cholesterol.

In the lower branch, 45.161% of the female who have a normal cholesterol and are drinking alcohol, but engaged in sports, have low blood pressure. Age is effective in this branch. Thus, 50% of them younger than 62,5 have low blood pressure.

43.62% of the male who has normal cholesterol, not using alcohol, and are engaging in sports have low blood pressure. 65.079% of the women who have normal cholesterol are drinking alcohol, and who are under the 67,5 age have a normal cholesterol. 96% of the male who have normal cholesterol, drinking alcohol, and are engaging in sports have low blood pressure.

Conclusion

This study attempts to present a data analytics in health care data. Preventive health care practices aim to implement customer value maximization manner. The results show that the proposed method could identify and evaluate critical points. A practical application of the method is also presented. Here, the groups which is at most risk are identified and the risk level for each category is determined. Health systems are facing a number of challenges in the cost-effective delivery of health care with aging populations and a number of diseases such as obesity, cancer, and diabetes increasing in prevalence. At the same time, the life sciences industry is also faced with historically low productivity. However, this has emerged as a science that can help tackle these challenges. The move toward electronic medical records in health systems has provided a rich source of new data for conducting research into the pathophysiology of disease (Kalaitzopoulos et al., 2016). Data analytics in health care will help the patients to be able to get treated at the right time. Also, this will help improve medical outcomes while also reducing the cost associated with mistreatment or overtreatment.

References:

Adriaans & Zantinge (1997). Data Mining, Addison-Wesley, p.5.

Andrew Taylor (1998). TQM implementation: organisation practices and information infrastructures, Omega, Volume 26, Issue 6, December 1998, Pages 715-727, ISSN 0305-0483, http://dx.doi.org/10.1016/S0305-0483(98)00022-X.

Andrei et al. (2015). Assessment and Continuous Improvement of Information Security Based on TQM and Business Excellence Principles, Procedia Economics and Finance, Volume 32, 2015, Pages 352-359, ISSN 2212-5671.

Ayako Suzuki et al. (2015). Comedications alter drug-induced liver injury reporting frequency: Data mining in the WHO VigiBase™, Regulatory Toxicology and Pharmacology, Volume 72, Issue 3, August 2015, Pages 481-490, ISSN 0273-2300.

Chad-Ton et al. (2006). Data mining for the diagnosis of type II diabetes from three-dimensional body surface anthropometrical scanning data, Computers & Mathematics with Applications, Volume 51, Issues 6–7, March–April 2006, Pages 1075-1092.

Chao-Hui, Jessie & Vincent (2011). A novel data mining mechanism considering bio-signal and environmental data with applications on asthma monitoring, Computer Methods and Programs in Biomedicine, Volume 101, Issue 1, January 2011, Pages 44-61, ISSN 0169-2607.

Cheng-Ding, Wang, & Jiang (2011). Using data mining techniques for multi-diseases prediction modeling of hypertension and hyperlipidemia by common risk factors, Expert Systems with Applications, Volume 38, Issue 5, May 2011, Pages 5507-5513, ISSN 0957-4174.

Ching-Wu, Hong-Qing Yao & Wu (2013). Applying data mining techniques to analyze the causes of major occupational accidents in the petrochemical industry, Journal of Loss Prevention in the Process Industries, Volume 26, Issue 6, November 2013, Pages 1269-1278, ISSN 0950-4230.

Ching-Wu et al. (2012). Applying data mining techniques to explore factors contributing to occupational injuries in Taiwan's construction industry, Accident Analysis & Prevention, Volume 48, September 2012, Pages 214-222, ISSN 0001-4575.

Daghfous & Barkhi (2009). The strategic management of information technology in UAE hotels: An exploratory study of TQM, SCM, and CRM implementations, Technovation, Volume 29, Issue 9, September 2009, Pages 588-595, ISSN 0166-4972.

Dheeraj et al. (2015). Exploring factors associated with pressure ulcers: A data mining approach, International Journal of Nursing Studies, Volume 52, Issue 1, January 2015, Pages 102-111, ISSN 0020-7489.

Dimitris, Ketan & Erfan (2016). Chapter 2 - Advancements in Data Management and Data Mining Approaches, In Translational Medicine, edited by Aamir Shahzad, Academic Press, Boston, 2016, Pages 35-53, ISBN 9780128034606.

Duen-Yian, Ching-Hsue & Yen-Wen (2011). A predictive model for cerebrovascular disease using data mining, Expert Systems with Applications, Volume 38, Issue 7, July 2011, Pages 8970-8977, ISSN 0957-4174.

Fields (1980). Client tracking: The development of a state-wide hypertension information system:, Massachusetts Department of Public Health, Division of Preventive Medicine, Boston, Massachusetts 02111. Sharon Spaight, Preventive Medicine, Volume 9, Issue 3, May 1980, Pages 429-430, ISSN 0091-7435.

Freitas (2002). Data mining and Knowledge Discovery with Evolutionary Algorithms, Springer, p.1-13.

Giudici (2003). Applied Data Mining: Statistical Methods for Business and Industry, Wiley, p.1.

Hian & Gerald (2005). Data Mining Applications in Healthcare Journal of Healthcare Information Management, Vol. 19.

Han & Kamber (2006). Data Mining Concepts and Techniques, Morgan Kaufmann, 2006, p.285.

IBM SPPS Decision Trees 22, p.1.

Jinn-Yi, Tai-Hsi & Chuan-Wei (2011). Using data mining techniques to predict hospitalization of hemodialysis patients, Decision Support Systems, Volume 50, Issue 2, January 2011, Pages 439-448, ISSN 0167-9236.

Joaquim & Celeste (2012). Finding occupational accident patterns in the extractive industry using a systematic data mining approach, Reliability Engineering & System Safety, Volume 108, December 2012, Pages 108-122, ISSN 0951-8320.

Joo, & Wang (2006) Relationship between total quality management (TQM) and continuous improvement of international project management (CIIPM), Technovation, Volume 26, Issues 5–6, May–June 2006, Pages 716-722, ISSN 0166-4972.

Katzman & Paushter (2016). Building a culture of continuous quality improvement in an academic radiology department, Journal of the American College of Radiology, Available online 18 February 2016, ISSN 1546-144.

Ling Chen et al. (2016). Personal health indexing based on medical examinations: A data mining approach, Decision Support Systems, Volume 81, January 2016, Pages 54-65, ISSN 0167-9236.

Liu (2007). Web Data Mining, Springer, p.59-60.

Management Overview (2016). Retrieved from:https://docs.oracle.com/cd/B28359_01/datamine.111/b28129/process.htmhttp://asq.org/learn-about-quality/total-quality-management/overview/overview.html.

Martínez-Lorente, Sánchez-Rodríguez, & Dewhurst (2004). The effect of information technologies on TQM: An initial analysis, International Journal of Production Economics, Volume 89, Issue 1, 8 May 2004, Pages 77-93, ISSN 0925-5273.

Maryam & Behrouz (2014). A new scoring system for assessing the risk of occupational accidents: A case study using data mining techniques with Iran's Ministry of Labor data, Journal of Loss Prevention in the Process Industries, Volume 32, November 2014, Pages 443-453, ISSN 0950-4230.

Muthu et al. (2012). Data mining technique for automated diagnosis of glaucoma using higher order spectra and wavelet energy features, Knowledge-Based Systems, Volume 33, September 2012, Pages 73-82, ISSN 0950-705.

Nenonen (2013). Analysing factors related to slipping, stumbling, and falling accidents at work: Application of data mining methods to finnish occupational accidents and diseases statistics database, Applied Ergonomics, Volume 44, Issue 2, March 2013, Pages 215-224, ISSN 0003-6870.

Nicole et al. (2010). Sources of health information related to preventive health behaviors in a national study, American Journal of Preventive Medicine, Volume 38, Issue 6, June 2010, Pages 620-627.e2, ISSN 0749-3797.

Preventive Health Care (2016). Retrieved from:https://en.wikipedia.org/wiki/Preventive_healthcare.

Steele & Schomer (2009). Continuous quality improvement programs provide new opportunities to drive value innovation initiatives in hospital-based radiology practices, Journal of the American College of Radiology, Volume 6, Issue 7, July 2009, Pages 491-499, ISSN 1546-1440.
Seyed et al. (2014). Identifying high-cost patients using data mining techniques and a small set of non-trivial attributes, Computers in Biology and Medicine, Volume 53, 1 October 2014, Pages 9-18, ISSN 0010-4825.
Shih et al. (2011). Health information systems in small practices: improving the delivery of clinical preventive services, American Journal of Preventive Medicine, Volume 41, Issue 6, December 2011, Pages 603-609, ISSN 0749-3797.
Taylor & Wright (2006). The contribution of measurement and information infrastructure to TQM success, Omega, Volume 34, Issue 4, August 2006, Pages 372-384, ISSN 0305-0483, http://dx.doi.org/10.1016/j.omega.2004.12.003.
Tan, Steinbach & Kumar (2006). Introduction to Data Mining, p. 151-152.
Tzeng-Ji, Li-Fang & Shinn-Jang (2003). Application of a data-mining technique to analyze coprescription patterns for antacids in Taiwan, Clinical Therapeutics, Volume 25, Issue 9, September 2003, Pages 2453-2463, ISSN 0149-291.
Valmohammadi & Roshanzamir (2015). The guidelines of improvement: Relations among organizational culture, TQM and performance, International Journal of Production Economics, Volume 164, June 2015, Pages 167-178, ISSN 0925-5273.
Vijayakrishnan et al. (2014). Prevalence of heart failure signs and symptoms in a large primary care population identified through the use of text and data mining of the electronic health record, Journal of Cardiac Failure, Volume 20, Issue 7, July 2014, Pages 459-464, ISSN 1071-9164.
What is Data Mining? (2016). Retrieved from: 24.03.2016. https://docs.oracle.com/cd/B28359_01/datamine.111/b28129/process.